

Canonical Frames for Planar Object Recognition*

Charles A. Rothwell¹, Andrew Zisserman¹, David A. Forsyth² and Joseph L. Mundy³

¹ Robotics Research Group, Department of Engineering Science, Oxford University, England.

² Department of Computer Science, University of Iowa, Iowa, USA.

³ The General Electric Corporate Research and Development Laboratory, Schenectady, NY, USA.

Abstract. We present a canonical frame construction for determining projectively invariant indexing functions for non-algebraic smooth plane curves. These invariants are semi-local rather than global, which promotes tolerance to occlusion.

Two applications are demonstrated. Firstly, we report preliminary work on building a model based recognition system for planar objects. We demonstrate that the invariant measures, derived from the canonical frame, provide sufficient discrimination between objects to be useful for recognition. Recognition is of partially occluded objects in cluttered scenes. Secondly, jigsaw puzzles are assembled and rendered from a single strongly perspective view of the separate pieces. Both applications require no camera calibration or pose information, and models are generated and verified directly from images.

1 Introduction

There has been considerable recent success in using projective invariants of plane algebraic curves as index functions for recognition in model based vision [10, 15, 23, 28]. Less attention has been given to invariants for smooth non-algebraic curves. In this paper we present a novel and simple method of constructing a family of invariants for non-convex smooth curves.

Lamdan *et al.* [18] proposed and implemented a canonical frame construction. We improve on this in two important ways:

1. The transformation here is *projective* not *affine*. Central projection between two planes is a type of projective transformation, not subject to the limitations on viewing distance required for affine approximation to hold. The affine transformation is only valid if the object depth variation is small compared to the camera viewing distance. Of course, projection includes the case that the transformation might actually be affine, because the affine group is a sub-group of the projective group.
2. Recognition is entirely via index functions based on projective invariants. In [18] recognition was a mixture of indexing and Hough style voting.

As has been argued elsewhere [10, 23], there is considerable benefit in using invariants to imaging transformations as *indexing functions* for generating recognition hypotheses. In particular, such functions only involve image measurements and avoid comparison

* CAR acknowledges the support of GE. AZ acknowledges the support of the SERC. DAF acknowledges the support of Magdalen College, Oxford and of GE. JLM acknowledges the support of the GE Coolidge Fellowship. The GE CRD laboratory is supported in part by the following: DARPA contract DACA-76-86-C-007, AFOSR contract F49620-89-C-003.

against each object in the model library. Complexity is then $O(i^k)$ rather than $O(\lambda i^k m^k)$ for the pose based comparison, where i is the number of image features, λ the number of models, m the number of features per model, and k the number of features needed to compute invariants, or determine transformations where invariants are not used⁴. Recognition hypotheses are verified in both cases by back-projection from models to images, and determining overlap of projected model with image curves.

This paper largely follows the path suggested by Lamdan *et al.* [18], where a very good discussion is given of reasonable requirements for curve representation in order to facilitate recognition tolerant to occlusion and clutter. Briefly, indices should be local and have some redundancy (i.e. several per outline), so if one index is occluded there is a good chance recognition can proceed on other visible parts; they should be stable, so small perturbations in the curve (due to image noise) do not cause large fluctuations in index value; and they should have sufficient discriminatory power over models in the library (so all models do not have similar index values). All of these requirements are satisfied by the bitangent construction described here. Provided the object outline is sufficiently rich in structure there will be several such constructions for each object, and thus redundancy in the representation giving partial immunity to occlusion.

We briefly review previous methods for curve recognition under distorting imaging transformations in section 2. The canonical frame construction and invariant measures are described in section 3. We apply these techniques to model based recognition for a library of planar objects of arbitrary (but non-convex) shape. They are recognised from single perspective views (no affine approximation is assumed) in scenes in which there may be partial occlusion by other known objects, or unknown clutter. The process does not require camera calibration. This is described in section 4. Finally, in section 5 we show how these measures can be used to reassemble a jigsaw.

2 Background

The recognition of silhouettes of planar objects under 2D similarity transformation (plane rotation, translation and isotropic scaling) and affine transformation has been extensively studied. The curve differential invariant *curvature* κ and the (integral) invariant s (arc length) have played a significant role because they are clearly unaffected by the action of the plane Euclidean group. Matching of these invariant curve “signatures” $\kappa(s)$ or their integral $\theta(s)$ (where $\kappa = \theta'$) are routine in the vision literature [1]. Unfortunately, such *differential invariants* for projection (called Wilczynski’s invariants [19, 29]) require 7th order derivatives. This is clearly numerically infeasible. Even affine projection requires 5th order derivatives. Such high order derivatives are required to give invariance to both projection and reparameterisation⁵.

Given the impracticality of using differential invariants directly, a number of methods have been derived for matching smooth curves despite affine or projective distortion:

1. Semi-differential invariants:

An ingenious method, proposed and implemented independently by Van Gool *et*

⁴ This complexity analysis is not for the asymptotic case as we assume that the library, implemented as a hash table, is sparse. Should this not be the case we increase the dimension of the library by using further invariants.

⁵ Recall that the familiar Euclidean curvature has this invariance: $\kappa = (\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t))/(\dot{x}(t)^2 + \dot{y}(t)^2)^{3/2}$ irrespective of the parameterisation t . That is, t can be replaced by $f(t)$ without affecting the value of κ .

al. [27] and Barrett *et al.* [3], is to trade derivatives at a point for more points. They demonstrate that at a combinatorial cost (some “reference” points must be matched) projective differential invariants can be derived requiring only first or second derivatives.

2. Representation by algebraic curves:

Since invariants for algebraic curves are so well established it is natural to try and exploit them by “attaching” algebraic curves to smooth curves. The algebraic invariants of these attached curves are then used to characterise the non-algebraic curve. This is the approach taken in [12] for affine invariance and [9, 17] for projective invariance. The problem here is that such methods tend to be global. Consequently, the associated algebraic curves and their invariants change if part of the curve is occluded.

3. Distinguished points:

A common method is to determine distinguished points on the curve, such as inflections and corners, which can be located before and after projection. Such points then effectively represent the curve - either to determine the transformation (e.g. alignment[14]) or to form algebraic invariants. The disadvantage is curve information between these points is effectively wasted.

4. Distinguished frame:

The goal is to get to some distinguished frame from any starting point; usually the frame corresponding to the plane of the object. A typical method is to maximise a function over all possible transformations - the transformed frame producing the function maximum determines the distinguished frame. Brady and Yuille considered a function measuring compactness over orthography [7, 11]; Witkin and others texture isotropy (over orthography) [5, 16, 30]; Marinos and Blake texture homogeneity (over perspectivities) [20]; and more recently Blake and Sinclair [6] with compactness over projectivities. Once in the distinguished frame any measurements act as invariants (because the measurements are independent of the original frame and transformation). Again this is a global approach and degrades with occlusion. There are also problems of uniqueness if the cost function is not convex, i.e. there are many local maxima.

5. Canonical frame:

Distinguished points are used to transform a portion of the object curve to a canonical frame [18]. As for the distinguished frame, any measurement made in this frame is an invariant. However, the canonical frame does not carry over the disadvantages: i) it is semi-local (depends on more than a single point) but is not global; ii) the transformation to the canonical frame is unique.

3 Canonical Frame Construction

3.1 Projective Transformations

A projective transformation between two planes is represented as a 3×3 matrix acting on homogeneous coordinates of the plane. The homogeneous representation means that only ratios of matrix elements are significant, and consequently the transformation has 8 degrees of freedom. This transformation models the *composed* effects of 3D rigid rotation and translation of the world plane (camera extrinsic parameters), perspective projection to the image plane, and an affine transformation of the final image (which covers the effects of changing camera intrinsic parameters). Clearly, all of these separate

transformations cannot be uniquely recovered from the single 3×3 matrix, since there are 6 unknown pose parameters, and 4 unknown camera parameters (camera centre, focal length and aspect ratio). We therefore have 10 unknowns with 8 constraints.

The mapping of four points between the planes is sufficient to determine the transformation matrix \mathbf{T} (each point provides two constraints, therefore 4 independent points provide $4 \times 2 = 8$ constraints). Corresponding points (x_i, y_i) and (X_i, Y_i) are represented by homogeneous 3 vectors $(x_i, y_i, 1)^T$ and $(X_i, Y_i, 1)^T$. The projective transformation $\mathbf{x} = \mathbf{TX}$ is:

$$\begin{pmatrix} kx_i \\ ky_i \\ k \end{pmatrix} = \begin{pmatrix} T_a & T_b & T_c \\ T_d & T_e & T_f \\ T_g & T_h & 1 \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ 1 \end{pmatrix}$$

where k is an arbitrary non-zero scalar. Eliminating k gives eight simultaneous equations linear in the matrix elements:

$$\begin{aligned} x_i(T_g X_i + T_h Y_i + 1) &= T_a X_i + T_b Y_i + T_c \\ y_i(T_g X_i + T_h Y_i + 1) &= T_d X_i + T_e Y_i + T_f \end{aligned}$$

with $i \in \{1, \dots, 4\}$. These are straightforward to solve, for example by Gaussian elimination.

Projectivities form a group, so every action has an inverse, and the composition of two projectivities is also a projectivity. Consequently two images, from different viewpoints, of the same object are related by a projectivity. This result is used in the verification stage of matching.

3.2 Obtaining Four Distinguished Points

The aim here is to exploit a construction that is preserved under projection. Certain properties, such as tangency and point of tangency are preserved by projection [26]. We use tangency to select 4 distinguished points on the curve (see figure 1) and then determine the projection that maps these to the corners of a unit square in the *canonical frame*. This projectivity is then used to map the curve into this frame. Figure 2 demonstrates this process for one concavity of a spanner. The object curve, and any projective view of it, are mapped into the same curve. Consequently, any (metric) measurements made in this frame are invariant descriptors and hence may be used as index functions to recognise the object. For example the location of any point in the frame is an invariant; it is not necessary to use Euclidean such as curvature.

Lamdan *et al.* [18] used bitangents to obtain two of three points to define a canonical frame under affine transformations. The third point was obtained by introducing a line parallel to the bitangent line in contact with the apex of the concavity. Since parallelism is not preserved under projective transformations, we use tangency conditions to define our third and fourth points. The selection of the corners of a unit squares as the corresponding points in the canonical frame is arbitrary - any four points, no three of which are collinear will do.

Alternative constructions are possible using other projectively preserved properties. For example, inflections can be used in two ways: i) to define a distinguished point on the curve; and ii) to define a line which is tangent at the inflection (3 point contact with the curve). If a concavity contains an inflection (and therefore it will necessarily have at least two inflections), then the bitangent contact points and inflections can be used as the four correspondence points. We believe a construction based on inflections will not be

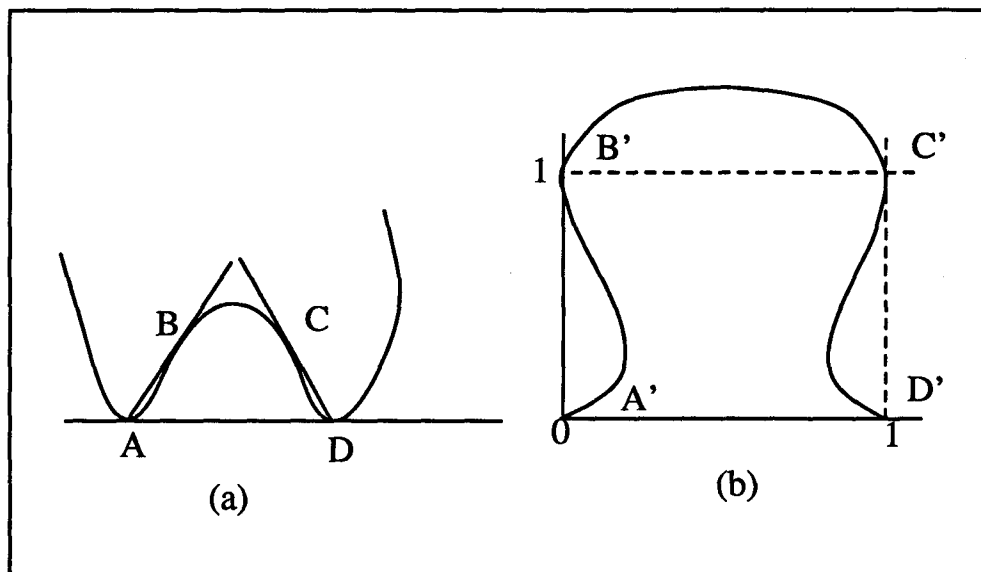


Fig. 1. (a) Construction of the four points necessary to define the canonical frame for a concavity. The first two points ($A D$) are points of bitangency that mark the entrance to the concavity. Two further distinguished points, ($B C$), are obtained from rays cast from the bitangent contact points and tangent to the curve segment within the concavity. These four points are used to map the curve to the canonical frame. (b) Curve in canonical frame. A projection is constructed that transforms the four points in (a) to the corner of the unit square. The same projection transforms the curve into this frame.

as stable (i.e. immune to small curve perturbations) as one based on tangencies, though we have not confirmed this⁶.

For this construction to be useful in model based vision it must satisfy two *sensible* and *useful* criteria:

1. Curves in the canonical frame for differing views of the same object should be very "similar".
2. Curves in the canonical frame from different objects should "differ" from each other.

The measures used to distinguish canonical curves are discussed in section 3.4.

3.3 Stability Over Views

The stability of the canonical frame representation is illustrated by figures 3a-d, which show three different views of the same spanner with extracted concavity and four reference points. The marked edge data is then mapped into the canonical frame. The curves in the canonical frame are almost identical. Representative images of other objects, a second

⁶ One interesting case of a further application of the bitangent construction is in forming invariants for curves with double points. This construction uses the dual space representation of a curve where the curve tangents (which are lines) are represented as homogeneous points in the plane. Then, a double point maps to a bitangent in the dual space of the curve (tangent space), and so invariants can be formed in the dual space.

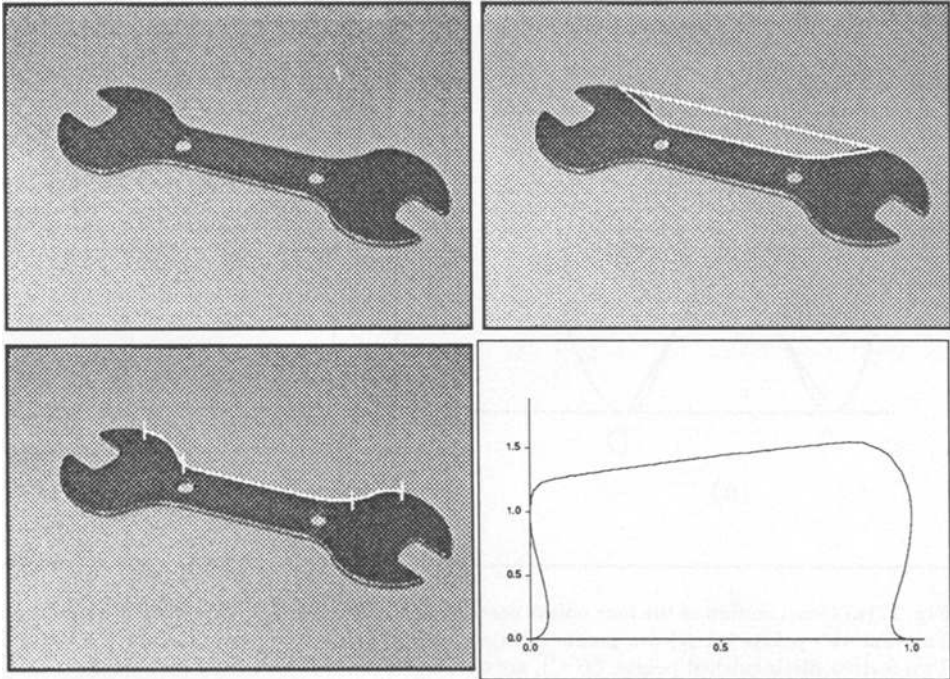


Fig. 2. Canonical frame transformation for a spanner concavity. (a) Original image. (b) Bitangent and tangents (see figure 1). These are determined via a bitangent detectors acting on image edge data. (c) Four distinguished points and concavity curve. (d) Projected curve in canonical frame. The curve passes through the corners of the unit square which are the projections of the four distinguished points. Note, the spanner has four external bitangents, four internal bitangents, and also bitangents which cross the boundary. Each of these can generate a curve in the canonical frame. Consequently, considerable redundancy is possible in the representation

spanner and a pair of scissors, are shown in figures 4 and 5 with their corresponding canonical curves (again from three views). Note that the jagged portion (A), of the curve in figure 5, varies over viewing position, but that the smooth portions (B), are consistent for all views. This is because (B) is produced by the plastic handles of the scissors, which are coplanar with the four reference points. The metal hinge is not coplanar with the reference point and so (A) is not positioned in a projectively invariant manner. This part of the curve must be excised. The variation emphasises the fact that the canonical frame construction is defined only on planar structures.

3.4 Index Functions and Discrimination

Since any measurements made in the canonical frame are invariant signatures for the curve, the question is what is the optimum set for discrimination over objects in the library? Clear criteria are that the number of measures should be reasonably small, but that there should be enough to discriminate objects from clutter, and that each one should be useful.

It appears that the most naive measurements, area moments, are stable and efficient discriminators. We use the area bounded by the x-axis and the curve. The moments

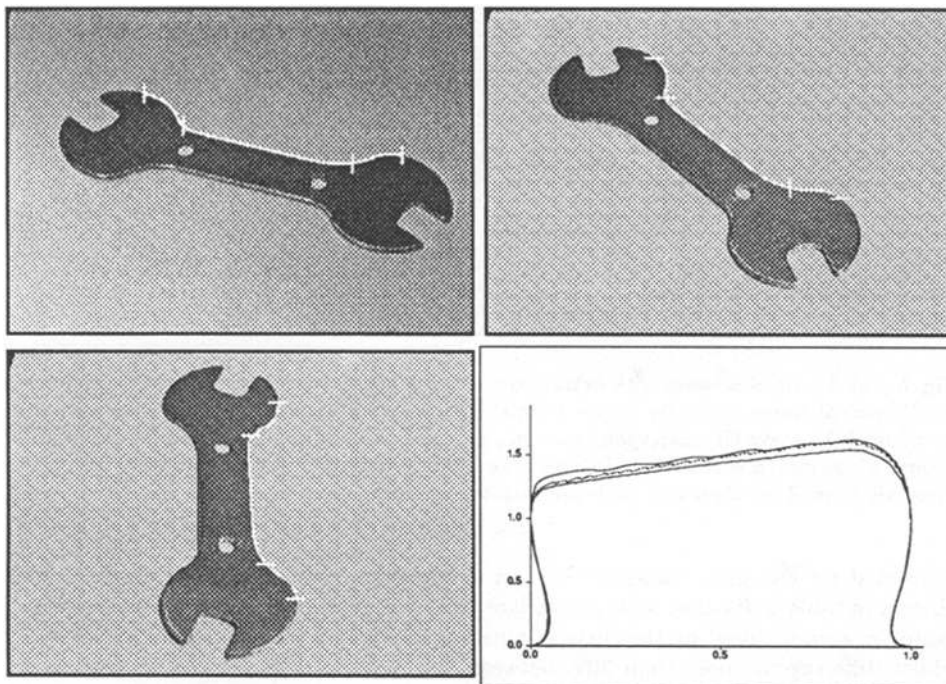


Fig. 3. (a) - (c) Three views of a spanner with extracted concavity curves and distinguished points marked. Note the very different appearance due to perspective effects. (d) Canonical frame curves for the three different views of the spanner. The curves are almost identical demonstrating the stability of the method. Of course the same curve would result from a projective transformation between the *object* and canonical frame.

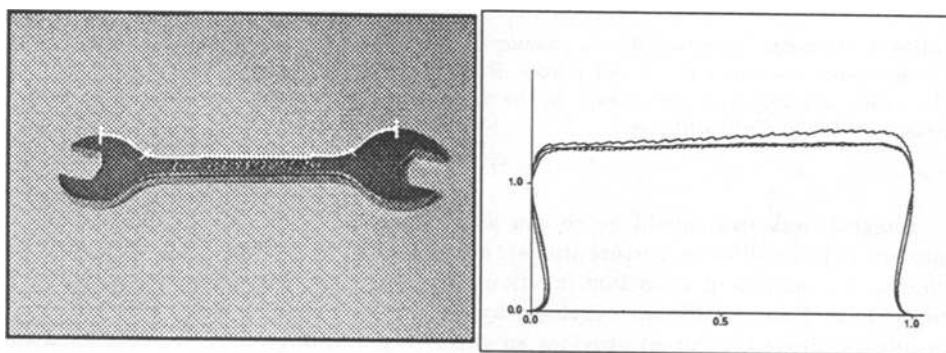


Fig. 4. (a) A second spanner with extracted concavity curves and distinguished points marked. (b) Canonical frame curves for this image and the same spanner from two other viewpoints. Again, the curves are very similar.

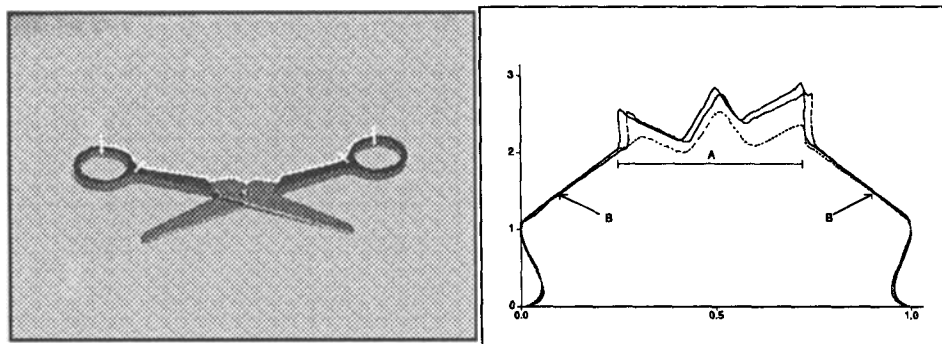


Fig. 5. (a) A pair of scissors with extracted concavity curves and distinguished points marked. (b) Canonical frame concavity curves from three views of the pair of scissors. The smooth end portions of the curve (B) correspond to regions of the concavity coplanar with the four reference points. These match well between images. The jagged portions (A) do not match as well because these are formed by edges non-coplanar with the reference points.

computed for the three views of the first spanner are given in table 1, and for all the objects in table 2. It is clear that in practice this construction gives very good results. For example, area enclosed by the curve is constant to 5% over views (viewpoint invariance), whilst differing by more than 30% between the spanner and scissors (discrimination). However, area alone could not reliably distinguish the two spanners (their areas differ by only 7%).

| view | Area | M_x | M_x^2 | M_y | M_y^2 |
|------|------|-------|---------|-------|---------|
| 1 | 1.35 | 0.516 | 0.341 | 0.720 | 0.686 |
| 2 | 1.40 | 0.518 | 0.343 | 0.743 | 0.732 |
| 3 | 1.42 | 0.516 | 0.341 | 0.756 | 0.756 |

Table 1. Moments computed for the spanner concavity when in the canonical frame (figure 3). The moments are about the x and y axes. Both the first and second moments are computed. The values are constant over change in the viewing position and so can be used as invariant measures to index into a library.

Integral measures should be chosen as they promote stability by smoothing noise (immunity to small curve perturbations). For example, measurement of curve arc length, which is the integral of a constant function along the curve, degrades systematically with image noise. However, the area enclosed by the curve (which may also be determined via an integral along the curve) provides an effective smoothing function, since local curve fluctuations have little effect on the total area enclosed by the curve.

To date, we have not investigated a set of 'optimal' measures that will provide maximum discrimination between objects with good stability. Future work will investigate two areas:

1. Given the model base we may perform a principal axis analysis to determine the dominant features (ie. the image shapes corresponding to the largest eigenvalues of

| view | Area | Mx | Mx^2 | My | My^2 |
|-----------|------|-------|--------|-------|--------|
| spanner 1 | 1.35 | 0.516 | 0.341 | 0.720 | 0.686 |
| scissors | 1.99 | 0.506 | 0.318 | 1.107 | 1.694 |
| spanner 2 | 1.26 | 0.507 | 0.334 | 0.665 | 0.584 |

Table 2. Moments for the three different objects are sufficiently different that they can be used for model discrimination. Note for example, that the measures for Mx^2 appear to be very similar, but because the values of Mx^2 are very stable the small differences provide sufficient discrimination.

the library covariant matrix) that provide the best discrimination between different models. One obvious problem with this approach is that the eigenvectors will be of high dimension, and so may not be realistically computable.

2. We may transform the data in the canonical frame to a set of orthogonal functions, for example using a Walsh or cosine transform, and use the transform coefficients as indexing values.

We are currently investigating different choices of index function, but in the demonstration of object recognition given in the next section we simply use area based measurements.

4 Model Based Recognition

This closely follows the system described in [23] where more details are given. There are two stages:

Pre-processing:

1. Model Acquisition:

Models are extracted directly from a single image of the unoccluded object. The edgel list is stored for later use in the verification process. Segmentation is carried out as described below to delimit concavities. No measurements are needed on the actual object, nor are pose or camera intrinsic parameters required.

2. Add to model library:

Invariant vectors of measures are calculated as described in section 3.4. Each component of this vector is an invariant measure that may be used as an index to the object. These vectors are entered into a library which will be accessed as a hash table.

Recognition:

1. Extract concavities:

Feature extraction and segmentation is carried out as below to delimit concavities.

2. Compute indices for each concavity:

As described in section 3.4.

3. Index into library:

If the index key corresponds to a table entry this is used to generate a recognition hypothesis.

4. Hypothesis Verification:

Verification proceeds in two phases (both based on the verification procedure of [23]):

- Check that the measured and expected model curves in the canonical frame are similar (that is lie close to each other).
- Project the edgel data from an acquisition image onto the current image. If sufficient projected edges overlap the target image edgels, the match is accepted. Note that the projective transformation between acquisition image and target image is computed directly from the correspondence of the four points used in the canonical frame construction.

4.1 Segmentation

A local implementation of Canny's edge detector [8] is used to find edges to sub-pixel accuracy. These edge chains are linked, extrapolating over any small gaps. Concavities are detected by finding bitangent lines.

Bitangent lines are found by computing approximations to the tangents of the curve, and representing these as points in the space of lines on the plane. Pairs of points that are close together in this space are found by a coarse search. These pairs represent approximate bitangents, which are refined through a convex hull construction.

4.2 Experimental Results

Some initial results of the system are demonstrated in figures 6 and 7. At present the model base consists of 5 objects: 2 spanners, scissors, pliers and a hacksaw. The figures demonstrate recognition under perspective of two models from this library despite the presence of partial occlusion and other objects (clutter) not in the library. Note there is a two fold ambiguity in the matching of curve tangent points to canonical frame. The matching depends solely on the ordering around the curve. To overcome this problem indexes and curves for both orderings are stored. Any problems with local symmetry in the concavity giving rise to ambiguous matches will be detected by back projection during the verification process.

5 Solving Jigsaw Puzzles

We have selected the problem of assembling a jigsaw puzzle to illustrate the shape discriminating power of the canonical frame approach. Jigsaw assembly has also long been considered a challenging vision task [13, 31]. The idea is that matching pieces will have the same invariant signatures for the tab and slot curves. We image a jumble of (unoccluded) puzzle pieces under significant perspective distortion and then "assemble" the puzzle by matching the pieces using canonical frame matching. The pieces are assumed to be planar but not necessarily coplanar with each other (in practice the pieces are not even taken from a single image). The assembly process is carried out by mapping the pieces to a common canonical frame and then aligning the matching curves. The texture patterns on the pieces are not used in the matching process but we warp the image of each piece to portray the assembly in a single plane. This experiment illustrates that the invariants calculated from the canonical frame can be used to compare unknown objects in a single image as well as classify objects from a library of model curves. Such comparison tasks are not readily tackled with conventional model-based recognition systems.

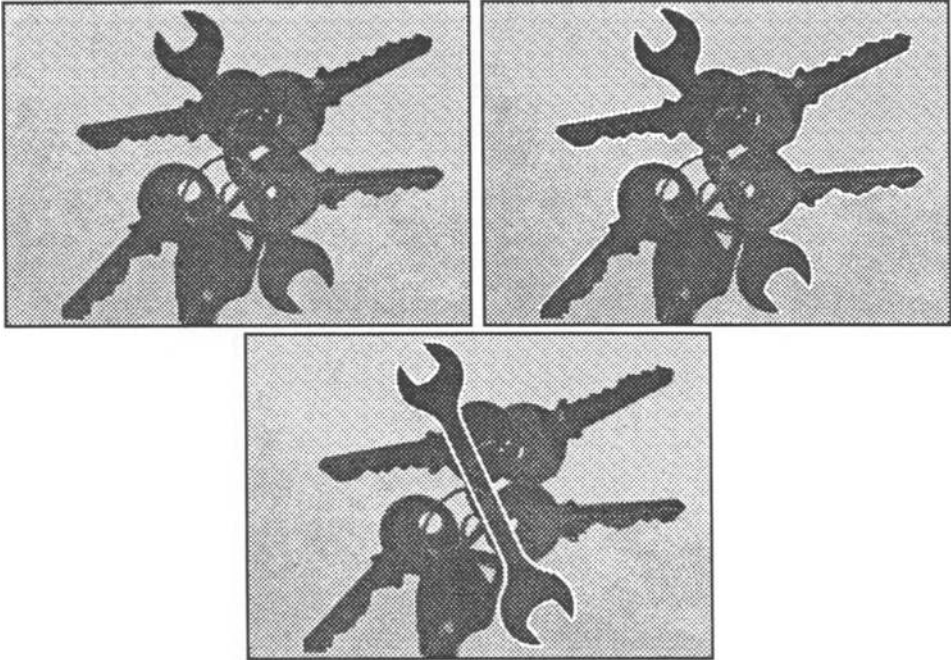


Fig. 6. (a) Spanner almost entirely occluded by keys. The keys are not the library, and are clutter in this scene. (b) Detected concavities, highlighted in white, which are used to compute indexes (c) The spanner which is the only model in the scene contained in the library, is recognised from the end slot concavity. The projected outline used for verification is highlighted in white.

5.1 Matching details

Edge pieces are extracted from unoccluded views of the pieces using a Canny [8] edge detector. Each piece is assumed to have four sides and to only connect to at most four other pieces⁷. Each of the four sides of a piece therefore either represent individual shape descriptors that match other pieces within the jigsaw or are edge pieces. Each of the side curves is then classified as either a straight side piece, a tab, or a slot, depending on the general shape of the curve in relation to the rest of the piece of which it is part. Each curve classed as a tab or a slot contains at least one significant concavity that corresponds to the tab or slot. We map⁴ each of the concavities into the canonical frame and search for the unique matching side. Once this is found the pieces can be joined together.

5.2 Reconstruction

The first corner piece found is used as the bottom left hand piece of the completed puzzle (a corner piece has two straight sides). This piece is used as the base unit square in the canonical frame, on which the puzzle is built. The piece immediately to its right can be mapped into the corner piece's image frame using the image to canonical frame mappings of the interlocking tabs and slots. Once this transformation is known the grey level values

⁷ This restriction is for implementation purposes only, and does not reduce the value of the demonstration.

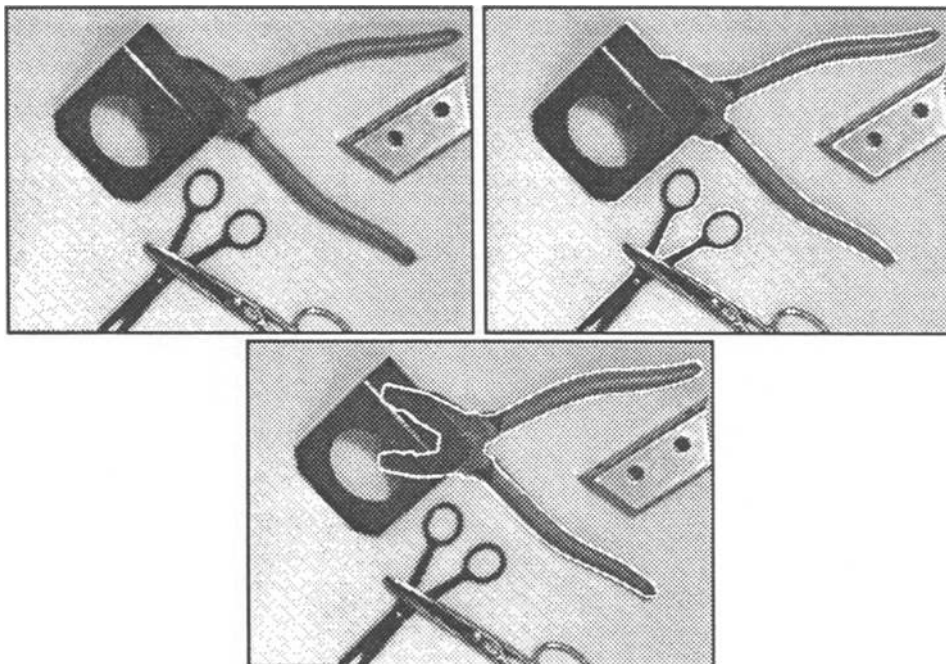


Fig. 7. (a) Image of various planar objects. (b) Concavities, highlighted in white, which are used to compute indexes (c) The pliers which are the only model in the scene contained in the library, is recognised and verified by projecting the edgels from an acquisition image, and checking overlap with edgels in this image.

of pixels within the bounds of the new piece can be projected into the frame of the corner piece, and so effect the joining of the two pieces. A similar process can be performed to map the piece directly above the corner piece into the corner image frame.

The piece diagonally above and to the right of the corner interlocks both the pieces above and to the right of the corner. We therefore do a least squares fit to determine the projectivity from the eight correspondences, and again render by projecting image grey values. A similar process is applied to the rest of the pieces.

Two examples of this are shown in figures 8 and 9. Both show the original pieces, and the final assembled and rendered puzzle.

This is a $O(n^2)$ algorithm (with n the number of pieces). Extracting indexes in the canonical frame, building a hash table, and using these indices for matching as in the recognition system, would reduce the complexity to $O(n)$, but here n is small and the time taken in computing matches negligible.

Some extensions are obvious:

1. The final reconstruction in the canonical frame should be mapped to a rectangle with the correct aspect ratio for the assembled jigsaw to remove any projective effects, or at least to a frame in which corners have right angles.
2. There are gaps in the assembled pattern arising because jigsaw pieces are mapped by a transformation determined by only a small part of the outline. This can be improved by determining the projectivity from all distinguished points around each piece using least squares.

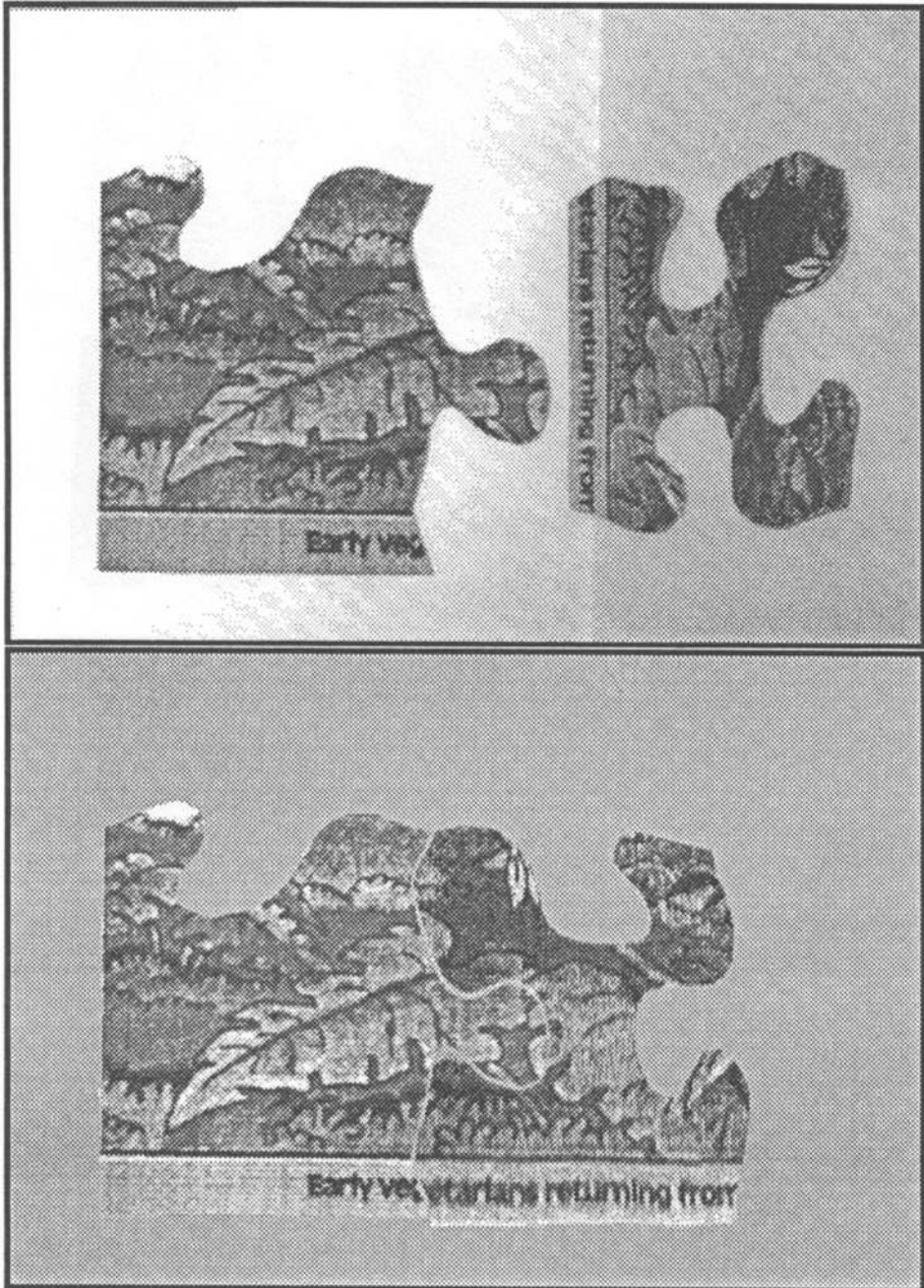


Fig. 8. (a) Two pieces of a jigsaw, with (b) the assembled and rendered solution. The puzzle is solved and rendered using only information from this image. No camera intrinsic parameters or pose information is needed. Note the large perspective distortion of the pieces in the original image (a) which are not in the same plane (the right hand piece lies in a plane at about 45° to the plane of the other piece).

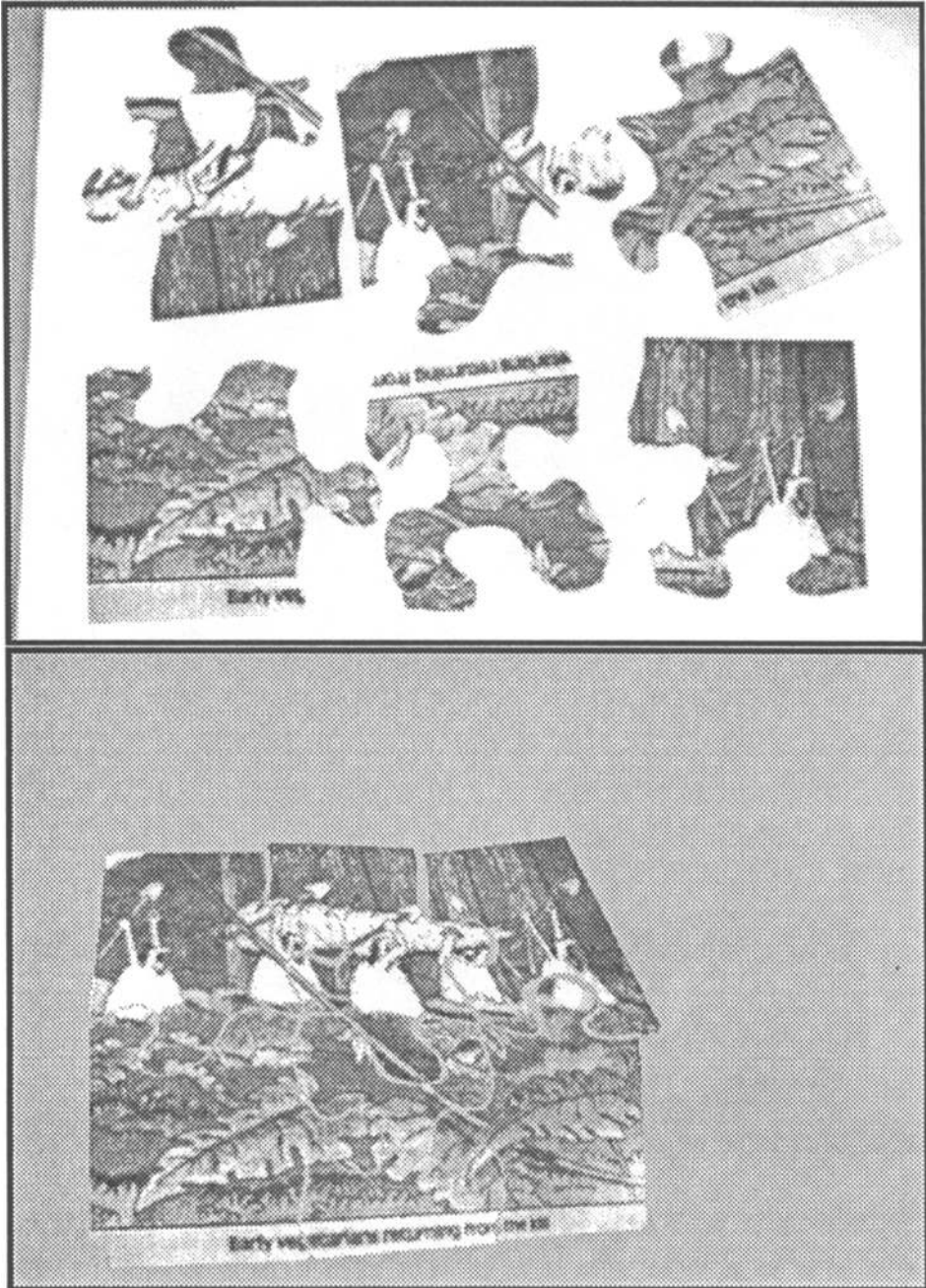


Fig. 9. (a) A six piece jigsaw, with (b) the assembled and rendered solution. The puzzle is solved and rendered using only information from this image. No camera intrinsic parameters or pose information is needed.

3. Other constraints - such as collinearity of the outer boundaries - could be incorporated by iteratively minimising a cost function.

6 Conclusions and Future Work

We have demonstrated recognition of non-algebraic planar objects from perspective images. The work is currently being extended in a number of ways:

1. At present there are five objects in the library. We are currently including more objects. Efficient development will require attention to the measures used for index functions.
2. We have assumed that the uncalibrated imaging process may be modeled by a projectivity. This is exact for a pin-hole camera, but corrections must be applied if radial-distortion is present. We are currently evaluating this correction [4, 25].
3. We can observe the reliability of an invariant measure by perturbing the distinguished points and recomputing the invariant values. This will be of benefit both during library construction and during the recognition process: i) We only use invariant indexes that are affected little by the bitangent locations, and ii) during verification confidence in a match is weighted by the stability of the invariant measure.
4. At present we use only concavities (exterior bitangents). This does not exploit the full structure of the curve. We wish to limit the bitangents used to those that do not cross the curve, but this does not prevent the use of internal bitangents. Using these will further improve immunity to occlusion.
5. There are obvious extensions for computing canonical frames for non-smooth curves. If a tangency discontinuity is observed we can use the two tangents immediately either side of the discontinuity as reference lines. We then find two more points or lines and uniquely determine the map to the canonical frame.

Acknowledgements

We are very grateful to both Professor Christopher Longuet-Higgins and Margaret Fleck for suggesting the jigsaw puzzle problem. We are also grateful to Mike Brady and Andrew Blake for useful discussions.

References

1. Asada, H. and Brady, M. "The Curvature Primal Sketch," *PAMI-8*, No. 1, p.2-14, January 1986.
2. Ayache, N. and Faugeras, O.D. "HYPER: A New Approach for the Recognition and Positioning of Two-Dimensional Objects," *PAMI-8*, No. 1, p.44-54, January 1986.
3. Barrett, E.B., Payton, P.M. and Brill, M.H. "Contributions to the Theory of Projective Invariants for Curves in Two and Three Dimensions," Proceedings First DARPA-ESPRIT Workshop on Invariance, p.387-425, March 1991.
4. Beardsley, P.A. "Correction of radial distortion," OUEL internal report 1896/91, 1991.
5. Blake, A. and Marinos, C. "Shape from Texture: Estimation, Isotropy and Moments," *Artificial Intelligence*, Vol. 45, p.232-380, 1990.
6. Blake, A. and Sinclair, D. On the projective normalisation of planar shape, *TR OUEL*, in preparation, 1992.
7. Brady, M. and Yuille, A. "An Extremum Principle for Shape from Contour," *PAMI-6*, No. 3, p.288-301, May 1984.

8. Canny J.F. "A Computational Approach to Edge Detection," *PAMI-6*, No. 6. p.679-698, 1986.
9. Carlsson, S. "Projectively Invariant Decomposition of Planar Shapes," *Geometric Invariance in Computer Vision*, Mundy, J.L. and Zisserman, A., editors, MIT Press, 1992.
10. Clemens, D.T. and Jacobs, D.W. "Model Group Indexing for Recognition," *Proceedings CVPR*, p.4-9, 1991, and *PAMI-13*, No. 10, p.1007-1017, October 1991.
11. Duda, R.O. and Hart P.E. *Pattern Classification and Scene Analysis*, Wiley, 1973.
12. Forsyth, D.A., Mundy, J.L., Zisserman, A.P. and Brown, C.M. "Projectively Invariant Representations Using Implicit Algebraic Curves," *Proceedings ECCV1*, Springer-Verlag, p.427-436, 1990.
13. Freeman, H. and Gardner, I. "Apictorial Jigsaw Puzzles: The Computer Solution of a Problem in Pattern Recognition," *IEEE Transaction in Electronic Computers*, Vol. 13, No. 2, p.118-127, April 1964.
14. Huttenlocher, D.P. and Ullman, S. "Object Recognition Using Alignment," *Proceedings ICCV1*, p.102-111, 1987.
15. Huttenlocher D.P. "Fast Affine Point Matching: An Output-Sensitive Method," *Proceedings CVPR*, p.263-268, 1991.
16. Kanatani, K. and Chou, T-C. "Shape From Texture: General Principle," *Artificial Intelligence*, Vol. 38, No. 1, p.1-48, February 1989.
17. Kapur, D. and Mundy, J.L. "Fitting Affine Invariant Conics to Curves," *Geometric Invariance in Computer Vision*, Mundy, J.L. and Zisserman, A., editors, MIT Press, to appear 1992.
18. Lamdan, Y., Schwartz, J.T. and Wolfson, H.J. "Object Recognition by Affine Invariant Matching," *Proceedings CVPR*, p.335-344, 1988.
19. Lane, E.P. *A Treatise on Projective Differential Geometry*, University of Chicago press, 1941.
20. Marinos, C. and Blake, A. "Shape from Texture: the Homogeneity Hypothesis," *Proceedings ICCV3*, p.350-354, 1990.
21. Mohr, R. and Morin, L. "Relative Positioning from Geometric Invariants," *Proceedings CVPR*, p.139-144, 1991.
22. Mundy, J.L. and Heller, A.J. "The Evolution and Testing of a Model-Based Object Recognition System," *Proceedings ICCV3*, p.268-282, 1990.
23. Rothwell, C.A., Zisserman, A. Forsyth, D.A. and Mundy, J.L. "Fast Recognition using Algebraic Invariants," *Geometric Invariance in Computer Vision*, Mundy, J.L. and Zisserman, A., editors, MIT Press, 1992.
24. Semple, J.G. and Kneebone, G.T. *Algebraic Projective Geometry*, Oxford University Press, 1952.
25. Slama, C.C., editor, *Manual of Photogrammetry*, 4th edition, American Society of Photogrammetry, Falls Church VA, 1980.
26. Springer, C.E. *Geometry and Analysis of Projective Spaces*, Freeman, 1964.
27. Van Gool, L. Kempenaers, P. and Oosterlinck, A. "Recognition and Semi-Differential Invariants," *Proceedings CVPR*, p.454-460, 1991.
28. Wayner, P.C. "Efficiently Using Invariant Theory for Model-based Matching," *Proceedings CVPR*, p.473-478, 1991.
29. Weiss, I. "Projective Invariants of Shapes," *Proceeding DARPA Image Understanding Workshop*, p.1125-1134, April 1988.
30. Witkin, A.P., "Recovering Surface Shape and Orientation from Texture" *Artificial Intelligence*, 17, p.17-45, 1981.
31. Wolfson, H., Schonberg, E., Kalvin, A. and Lamdan, Y. "Solving Jigsaw Puzzles by Computer," *Annals of Operational Research*, Vol. 12, p.51-64, 1988.